Testing for Granger causality in the presence of measurement errors

Jonas Andersson *

September 24, 2004

Abstract

In this paper a potential problem with tests for Granger-causality is investigated. If one of the two variables under study, but not the other, is measured with error the consequence is that tests of forecastablity of the variable without measurement error by the variable with measurement error will be rejected less often than it should. Since this is not the case for the test of forecastability of the variable with measurement error by the one without there is a danger of concluding that one variable leads the other while it is in fact a feed-back relationship. The problem is illustrated by an example.

JEL: C32

Keywords: Granger causality, measurement error

1 Introduction

Since the ground-breaking work by Granger (1969), tests for what is now called Granger-causality have been employed to evaluate forecasting ability of one time series variable by another. Even though sometimes mixed up with the everyday-use word “causality” it can, at least rule out that one variable

*Department of Finance and Management, Norwegian School of Economics and Business Administration. The author thanks Loran Cholette and Jostein Lillestøl for valuable comments.

is causing another by the reasonable idea that for an event to cause another event it must at least precede it. Therefore it is perhaps as close as we can get in using data analysis to evaluate the philosophical concept of causality. Some variables in macroeconomics and finance are arguably measured with error. Examples are inflation, economic growth and volatility in financial markets. In the next section, a brief review of Granger causality and how to test it is given. Section 3 investigates properties of the test when one of the variables is measured with error. Section 5 concludes.

2 Granger causality

A time series variable x is said to fail to *Granger-cause* another variable y if the mean squared error (MSE) of a forecast of y_{t+s} based on $\mathcal{F}_t^{xy} = \{x_t, x_{t-1}, ..., y_t, y_{t-1}, ...\}$ is equal to the MSE of a foreacast based on $\mathcal{F}_t^{y} = \{y_t, y_{t-1}, ...\}$, $s > 0$. Tests of Granger-causality can e.g. be based on a vector autoregressive model, a multivariate MA-representation or a regression of y_{t+s} on $\mathcal{F}_t^{xy}$. See Hamilton (1994) for a review of such tests. For the purpose of this paper, the last of these approches is particularly helpful and therefore chosen. The test I consider is simply performed by testing the hypothesis

$$\begin{cases} H_0 : \alpha_1 = ... = \alpha_p = 0 \\ H_1 : \text{At least one } \alpha_j \neq 0 \end{cases} \tag{1}$$

where the parameters are given by the model

$$y_t = \alpha_0 + \alpha_1 x_{t-1} + ... + \alpha_p x_{t-p} + \beta_1 y_{t-1} + ... + \beta_p y_{t-p} + a_{1t} \tag{2}$$

where a_{1t} is a zero mean strict white noise. The choice of the lag length, p, is of great importance for this type of analysis but is not the object of this paper. Therefore, it is assumed to be known. The possibility that a_{1t} can be autocorrelated in practice is not considered either.

In order to test H_0 the model

$$y_t = \gamma_0 + \gamma_1 y_{t-1} + ... + \gamma_p y_{t-p} + a_{0t} \tag{3}$$

is estimated as well. We form

$$S_1 = \frac{T(RSS_0 - RSS_1)}{RSS_1} \tag{4}$$

where

$$RSS_i = \sum_{t=1}^{T} \hat{a}_{it}^2, \tag{5}$$

and $i = 0, 1$, are the residual sum of squares for the null and alternative hypothesis, respectively. Then, under the null hypothesis, S_1 is asymptotically $\chi^2(p)$-distributed.

3 Measurement error

It is common that tests of Granger-causality are used both to investigate whether x fails to Granger-cause y and vice versa. This can be made, e.g. in order to establish whether events connected with inflation are preceding events connected with consumer behaviour or whether the opposite is true.

Assume now that x is measured with error while y is not. Thus, x can be written

$$X_t = x_t + e_t \tag{6}$$

where X_t is the observed value of x and e_t is a measurement error which is assumed to be a strict white noise with variance σ_e^2. As an example we consider the case where $p = 1$ and $\alpha_0 = 0$. Then

$$\hat{\alpha}_1 \xrightarrow{p} \alpha_1 \frac{\sigma_{xy} - \sigma_x^2 \sigma_y^2}{\sigma_{xy} - (\sigma_x^2 + \sigma_e^2)\sigma_y^2} \leq \alpha_1 \tag{7}$$

showing that we will on average, underestimate the parameter α_1, representing the forecasting value of x on y.

If we instead test if y is useful in forecasting x, the measurement error ends up both in the dependent and independent variables. Maintaining that $p = 1$ and the absence of intercept the regression

$$y_t = \gamma_1 x_{t-1} + \delta_1 y_{t-1} + b_{1t} \tag{8}$$

is estimated. The probability limit of the OLS estimate of δ_1 is then

$$\hat{\delta}_1 \xrightarrow{p} \delta_1 \frac{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2}{(\sigma_x^2 + \sigma_e^2)\sigma_y^2 - \sigma_{xy}^2} + \gamma_1 \frac{\sigma_e^2 \sigma_{xy}}{(\sigma_x^2 + \sigma_e^2)\sigma_y^2 - \sigma_{xy}^2} \tag{9}$$

From (9) it can be seen that there is no clearcut inequality as in (7). Whether $\hat{\delta}_1$ converges to a quantity larger or smaller than δ_1 depends on the size and sign of σ_{xy}, the correlation between x and y, and the parameter γ_1.

Say, as an example, that $\sigma_{xy} > 0$ and $\gamma_1 > 0$. Then the inequality $plim\hat{\delta}_1 > \delta_1$ can occur.[1]. This is true for all values of σ_e^2 except the special case

$$\sigma_e^2 = \frac{\delta_1(\sigma_{xy}^2 - \sigma_x^2\sigma_y^2)}{\gamma_1\sigma_{xy}} \tag{10}$$

The equation (9) also indicates that we are dealing with, mainly, a small sample problem. The explanation to this is that, if $\delta_1 \neq 0$, the estimator $\hat{\delta}_1$ will converge, in probability, to a quantity not equal to zero and thereby cause a rejection of the null hypothesis that y is Granger causing x. In the next section the small sample problem is illustrated by means of a simulation study.

4 Simulation study

The calculations in Section 3 was instructive in order to see that the asymptotic consequence of measurement error in x was different for the test of forecasting power in x on y than for the test of forecasting power in y on x. However, it did not show how it affected the power of such tests. The question now is: Given that there is a feedback between x and y, does a measurement error in x cause more rejections in one of the tests than in the other?

In order to study this in the finite sample case data from a bivariate VAR(1)-models is generated.

$$\begin{cases} x_t = 0.5x_{t-1} + 0.2y_{t-1} + a_{x,t} \\ y_t = 0.2x_{t-1} + 0.5y_{t-1} + a_{y,t} \end{cases} \tag{11}$$

where $(a_{x,t}, a_{y,t})'$ is a normally distributed bivariate white noise with covariance matrix Σ_a, is considered. This example is a situation where there is a symmetry in the sense that y_t is Granger-caused by x_t "as much as" the opposite is true. However x_t is measured with error according to (6). The rejection of the two null hypotheses

$$H_{0xy} : x \text{ fails to Granger-cause } y$$

and

$$H_{0yx} : y \text{ fails to Granger-cause } x$$

[1] $plim$ indicates "limit in probability"

Given the data generating process above, both these hypotheses should, optimally, be rejected as often as possible.

The parameter of interest that I will vary is the covariance between $a_{x,t}$ and $a_{y,t}$. The results are presented in Table 1. The table shows empirical rejection rates when the nominal significance level is 5%. In the case of a

	$\Sigma_a = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\Sigma_a = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$	$\Sigma_a = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$
H_{0xy}	0.409	0.259	0.376
H_{0yx}	0.527	0.788	0.189

Table 1: Monte Carlo rejection rates (power) of the two tests H_{0xy} and H_{0yx} for three different Σ_a. The nominal significance level is 5%, the signal to noise ratio is one and the sample size, T is 100.

signal-to-noise ratio of one and a positive correlation between a_1 and a_2, as can be seen in Table 1, yield a power for the test of H_{0yx} which is substantially larger than for the test of H_{0xy}. The implication of this is that it is more likely that the conclusion is that y is driving x is more likely than the opposite. In the case of a negative correlation between a_1 and a_2, the opposite is true.

	$\Sigma_a = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\Sigma_a = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$	$\Sigma_a = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$
H_{0xy}	0.572	0.342	0.512
H_{0yx}	0.610	0.679	0.403

Table 2: Monte Carlo rejection rates (power) of the two tests H_{0xy} and H_{0yx} for three different Σ_a. The nominal significance level is 5%, the signal to noise ratio is 4 and the sample size, T is 100.

In Table 2 the case with a signal to noise ratio of 4 is considered. The tendency is the same while the relative decrease in measurement error is moving the power of the two tests closer.

5 Conclusion

The problem of measurement errors in one of the variables in tests of Granger-causality has been studied. In small samples where the correlation between the two variables are positive the problem occurs in that the variable measured with error is often mistakenly concluded to fail to Granger-cause the other variable while the Granger causality in the other direction is more often detected. This causes a tendency to conclude that one variable is driving the other while there is indeed a feedback relationship present.

References

C.W.J Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37:424–438, 1969.

J.D Hamilton. *Time Series Analysis*. Princeton, 1994.

Appendix

Consider the model

$$y_t = \alpha_1 x_{t-1} + \beta_1 y_{t-1} + a_{1t} \tag{12}$$

and

$$X_t = x_t + e_t. \tag{13}$$

where $E(x_t) = 0$. The OLS-estimator of α_1 can be written

$$\hat{\alpha}_1 = \frac{\sum_{t=1}^{T-1} y_t^2 \sum_{t=1}^{T-1} X_t y_{t+1} - \sum_{t=1}^{T-1} X_t y_t \sum_{t=1}^{T-1} y_t y_{t+1}}{\sum_{t=1}^{T-1} X_t^2 \sum_{t=1}^{T-1} y_t^2 - (\sum_{t=1}^{T-1} X_t y_t)^2} \tag{14}$$

Multiplying both the numerator and denumenator by $1/T^2$ and taking each of the terms in probability limit we obtain

$$\hat{\alpha}_1 \xrightarrow{p} \frac{\sigma_y^2(\alpha_1 \sigma_x^2 + \beta_1 \sigma_{xy}) - \sigma_{xy}(\alpha_1 \sigma_{xy} + \beta_1 \sigma_y^2)}{(\sigma_x^2 + \sigma_e^2)\sigma_y^2 - \sigma_{xy}^2} \tag{15}$$

which can be rewritten as (7).

Now keeping the variables X_{t-1} and y_{t-1} on the right-hand side of (12) but replacing the left-hand side with X_t we can write the OLS-estimator of δ_1 in the regression

$$y_t = \gamma_1 x_{t-1} + \delta_1 y_{t-1} + b_{1t} \tag{16}$$

as

$$\hat{\delta}_1 = \frac{\sum_{t=1}^{T-1} X_t^2 \sum_{t=1}^{T-1} y_t X_{t+1} - \sum_{t=1}^{T-1} X_t y_t \sum_{t=1}^{T-1} X_t X_{t+1}}{\sum_{t=1}^{T-1} X_t^2 \sum_{t=1}^{T-1} y_t^2 - (\sum_{t=1}^{T-1} X_t y_t)^2} \tag{17}$$

Again, multiplying both the numerator and denumenator by $1/T^2$ we obtain

$$\hat{\delta}_1 \xrightarrow{p} \frac{(\sigma_x^2 + \sigma_e^2)(\gamma_1 \sigma_{xy} + \delta_1 \sigma_y^2) - \sigma_{xy}(\gamma_1 \sigma_x^2 + \delta_1 \sigma_{xy})}{(\sigma_x^2 + \sigma_e^2)\sigma_y^2 - \sigma_{xy}^2} \tag{18}$$

which can be rewritten as (9).